

## **An opponent process for alcohol addiction based on changes in endocrine gland mass**

Omer Karin, Moriya Raz, Uri Alon\*

Dept. Molecular Cell Biology, Weizmann Institute of Science, Rehovot Israel 76100

\*corresponding author: [urialon@weizmann.ac.il](mailto:urialon@weizmann.ac.il)

### **Summary**

Consuming addictive drugs is often initially pleasurable, but escalating drug intake eventually recruits physiological “anti-reward” systems called opponent processes that cause tolerance and withdrawal symptoms. Opponent processes are fundamental for the addiction process, but their physiological basis is not fully characterized. Here, we propose an opponent processes mechanism centered on the endocrine stress-response, the HPA axis. We focus on alcohol addiction, where the HPA axis is activated and secretes  $\beta$ -endorphin, causing euphoria and analgesia. Using a mathematical model, we show that slow changes in HPA glands act as an opponent process for  $\beta$ -endorphin secretion. The model explains hormone dynamics in alcohol addiction, and experiments on alcohol preference in rodents. The opponent process is based on fold-change detection (FCD) where  $\beta$ -endorphin responses are relative rather than absolute; FCD confers vulnerability to addiction but has adaptive roles for learning. Our model suggests gland-mass changes as potential targets for intervention in addiction.

## **Introduction**

Drug addiction is a process in which the individual becomes increasingly occupied by drug-seeking and drug-taking behavior, escalates drug-taking over time, and has difficulties quitting. Addiction has several affective stages. While the contribution of each of these stages to the addiction process is debated (Wise and Koob, 2014), it is acknowledged that the stages are shared among many addictive drugs, including alcohol, opiates, and cocaine. The first stage is the initiation stage, which is related to positive reinforcement by drug administration. Next, the phenomenon of hedonic tolerance sets in: increasing amounts of drug are needed to produce the same effect. This leads to the maintenance stage in which the drug is taken in part to avoid the negative effects of withdrawal. Changes in learning and memory systems result in compulsive drug-seeking habits and increase subsequent risk of relapse (Hyman et al., 2006; Milton and Everitt, 2012; Robbins, 2002). Once drug-use is stopped, the withdrawal stage occurs, characterized by persistent negative affect (Koob, 2011; Koob and Le Moal, 2001). This provides negative reinforcement against cessation of drug use. Withdrawal gradually gives way to recovery (Figure 1A).

These affective changes have been studied by the opponent-process theory of addiction (Koob and Le Moal, 1997; Koob and Volkow, 2016; Solomon, 1980) (Figure 1ABC). This theory posits that chronic drug use activates anti-reward processes called opponent processes. Opponent processes are secondary slow processes that are activated by the drug and which antagonize the primary, pleasurable process (Fig 1A). After initial drug use, the opponent process causes a slight and slow-to-decay undershoot in mood. Repeated, extensive drug use sensitizes the secondary process and desensitizes the primary process (a phenomena known as “reward allostasis”), leading to deficient primary hedonic responses and exaggerated negative affect upon drug withdrawal. This explains the transition from the initial euphoric stage of drug use to the maintenance stage, where the drug is taken to avoid the negative effects of withdrawal that are due to the opponent process.

An important challenge for addiction research is to identify the molecular mechanisms of these opponent processes. The difficulty lies in the complexity of the physiological circuits that underlie addiction. These reward-processing circuits control pleasure, pain and reinforcement (Berridge and Robinson, 2016; Kelley and Berridge, 2002; Koob and Le Moal, 1997; Koob and Volkow, 2016; Solomon, 1980). Addictive drugs including alcohol, cocaine, and opiates affect the levels of specific neurotransmitters and other secreted molecules, which, in turn, affect the brain regions involved in reward processing (Koob and Volkow, 2016). Since these circuits are intricate and involve feedback over multiple timescales, mathematical models

are essential for analyzing their relation with addiction dynamics and for pointing to potential opponent processes.

Most proposals for opponent processes involve neurotransmitter circuits, primarily dopamine, a key neuromodulator of learning and motivation (Berridge and Robinson, 1998; Glimcher, 2011; Schultz, 1998; Schultz et al., 1997; Wise, 2004). One such model was developed by Gutkin et al. to study nicotine addiction (Gutkin et al., 2006). Gutkin et al. analyzed the interaction between nicotine and nicotine receptors expressed in neurons which secrete dopamine. They proposed a model where on the fast time scale nicotine causes dopamine to be secreted and upregulates the phasic dopamine response, while on a much slower timescale chronic nicotine administration activates an opponent process which downregulates tonic dopamine responses. Upon nicotine withdrawal, this slow process causes an undershoot in tonic dopamine. Other proposed opponent processes include the recruitment of stress-associated neurotransmitters such as dynorphin and CRH in the amygdala, and various other neurochemical systems (reviewed in (Koob and Volkow, 2016)).

Since opponent processes are important for understanding addiction, it is useful to propose additional physiological systems which can provide opponent process properties.

Here we propose a new class of opponent processes, based not on neurotransmitters but instead on endocrine circuits. For this purpose, we model the dynamics of an endocrine system which plays an important role in alcohol addiction: the secretion of endogenous opioids following HPA axis activation (Figure 1D). Endogenous opioids control subjective reward, euphoria and pain sensitivity (Darcq and Kieffer, 2018; Drews and Zimmer, 2010; Gerrits et al., 2003; Gianoulakis, 2004; Kiefer et al., 2002; Kuzmin et al., 1997; Mitchell et al., 2012; Roberts et al., 2000; Roth-Deri et al., 2008; Trigo et al., 2009; Van Ree, 1996). They are secreted in response to stress after activation of the HPA axis (Nikolarakis et al., 1986; Tsigos and Chrousos, 2002), and affect dopamine release (Spanagel et al., 1992). In particular,  $\beta$ -endorphin, which causes euphoria and analgesia by binding to the same receptor as morphine, has been tightly linked with addiction to alcohol and other substances which activate the HPA axis (Gianoulakis, 2004; Kiefer et al., 2002; Roth-Deri et al., 2008; Trigo et al., 2009). Rodents that lack  $\beta$ -endorphin or its receptor (mu-opioid receptor) show diminished alcohol self-administration (Hall et al., 2001; Racz et al., 2008; Roberts et al., 2000), and deficiency in  $\beta$ -endorphin during the weeks after alcohol withdrawal is associated with anxiety in patients (Kiefer et al., 2002).

The importance of  $\beta$ -endorphin for alcohol addiction raises the question of what its opponent process may be. One candidate is the secretion of the stress hormone cortisol from

the adrenal gland. Cortisol inhibits  $\beta$ -endorphin secretion both directly and by inhibition of CRH secretion. This was shown by experiments employing adrenalectomy, as well as by pharmacological manipulations such as the administration of dexamethasone and of the glucocorticoid synthesis inhibitor metyrapone (Guillemin et al., 1977; Nakao et al., 1978; Holaday et al., 1979; Lim et al., 1982; Rivier et al., 1982; Kreek et al., 1984; Hargreaves et al., 1987; Young, 1989). Cortisol is greatly elevated during periods of excessive alcohol consumption (Figure 1E, (Esel et al., 2001; Stalder et al., 2010)), and returns to baseline over a few weeks after alcohol withdrawal (von Bardeleben et al., 1989; Esel et al., 2001; Kiefer et al., 2002; Marchesi et al., 1997). Abolishing cortisol secretion by the removal of the adrenal glands diminishes alcohol preference in rodents. Alcohol preference in adrenalectomized animals is rescued by corticosterone replacement (Figure 1F) (Fahlke, 2000; Fahlke et al., 1994; Hansen et al., 1995; Lamblin and De Witte, 1996). Similar results were reported for cocaine (Goeders, 2002; Goeders and Guerin, 1996)). Since adrenalectomy removes the negative feedback from  $\beta$ -endorphin secretion (Young, 1989), this raises the possibility that cortisol secretion is an opponent process for  $\beta$ -endorphin. However, cortisol has a half-life of only 1 hour, similar to that of  $\beta$ -endorphin (Foley et al., 1979; McKay and Cidlowski, 2003). This is much faster than the timescale of days-weeks expected for an opponent process, suggesting that adrenal secretion of cortisol cannot by itself be the opponent process for  $\beta$ -endorphin secretion.

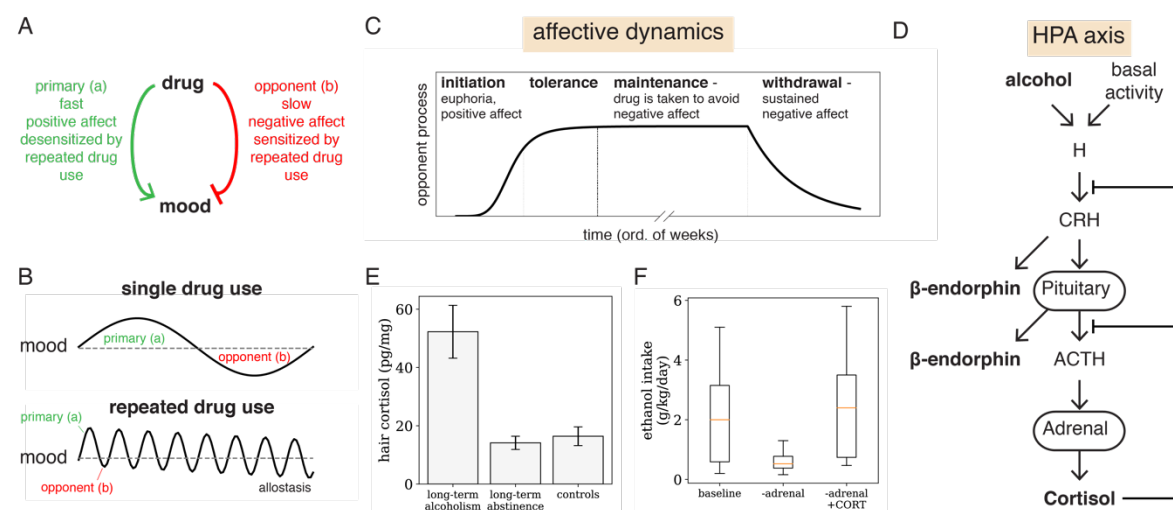
Here we propose that the opponent process for  $\beta$ -endorphin secretion may be due to structural changes of the HPA glands following repeated alcohol intake, namely the growth of the total mass of the hormone-secreting cells in the adrenal and pituitary glands. We show this by analyzing the dynamics of  $\beta$ -endorphin secretion following acute and repeated alcohol intake, using a recently-developed mathematical model of the HPA axis (Karin et al., 2020). The important aspect of this model is that it includes the changes in the total cell mass of the HPA glands due to cell proliferation and growth. This growth is caused by the HPA hormones, which act as growth factors for specific HPA glands. Changes in gland masses are well documented in conditions associated with HPA activation in humans (Amsterdam et al., 1987; Carey et al., 1984; Doppman et al., 1988; Ludescher et al., 2008; Nemeroff et al., 1992; Rubin et al., 1995, 1996), including enlarged adrenal glands in chronic alcohol abuse (Carsin-Vu et al., 2016). Adrenal weight in rodents also increases after alcohol administration (Adams and Hirst, 1984; Đikić et al., 2011; Mendelson et al., 1971).

The gland-growth mechanism adds a timescale of weeks to the hours-timescale of the HPA hormones. The model shows that while alcohol initially increases  $\beta$ -endorphin levels,

persistent HPA activation by alcohol causes the HPA glands to grow, causing increased cortisol secretion which renormalizes  $\beta$ -endorphin levels and suppresses them upon withdrawal. *The gland masses thus provide an opponent process*, explaining data on  $\beta$ -endorphin and HPA hormones during alcohol addiction and withdrawal, as well as the effect of adrenalectomy and cortisol replacement on alcohol preference in rodents. The main message of this study is thus that a weeks-scale feedback loop in which HPA gland masses changes over time provides an opponent process.

Using the model, we analyze the fundamental reason for the HPA opponent process: fold-change detection (FCD) for  $\beta$ -endorphin dynamics in response to drug inputs. FCD is a property of biological circuits (Adler and Alon, 2018; Shoval et al., 2010) where the output of the circuit responds to *relative* changes in the input, rather than absolute changes. The model makes several predictions that can be tested experimentally. We show, using a minimal model of reward optimization, that FCD circuits that control subjective reward are uniquely fragile to addiction, because they maintain sensitivity to reward at increasing levels of drug intake.

In addition to this fragility, FCD control of subjective reward has two beneficial functions for reward-based learning. The first is the ability to learn across many orders of magnitude of rewards. The second benefit is that FCD implements an important concept from reinforcement learning, called potential-based reward shaping (Ng et al., 1999), which helps an individual to learn by adding auxiliary rewards that guide exploration. Thus, FCD has beneficial properties for learning, while it has fragility to addiction.



**Figure 1. Opponent process theory explains the affective stages of addiction.** (A) According to the opponent process theory of drug addiction (Koob and Le Moal, 2001; Koob and Volkow, 2010), the intake of an addictive drug activates two processes: a fast, primary process that causes the pleasurable effects of the drug, and a slow “anti-reward” opponent process which results in negative affect. Repeated intake over long periods of time desensitizes the primary process and sensitizes the opponent process, resulting in allostasis of mood and sustained negative affect after drug withdrawal. (B) Schematic illustrations of the effects of single and repeated drug intake on mood. (C) The opponent

process framework can explain the stages of addiction to various drugs: initiation (weeks), tolerance that leads to maintenance (weeks to decades), and withdrawal (weeks to months). (D) Regulation of  $\beta$ -endorphin by alcohol.  $\beta$ -endorphin is secreted from POMC-neurons in response CRH from the hypothalamus, denoted H. CRH also causes pituitary corticotrophs to secrete ACTH, which is cleaved from the same peptide as  $\beta$ -endorphin. ACTH causes the adrenal gland to secrete cortisol, which, in turn, inhibits the secretion of CRH and the pituitary secretion of ACTH and  $\beta$ -endorphin. Cortisol is therefore a candidate component of an opponent process mechanism for the rewarding effects of  $\beta$ -endorphin. This is supported by data indicating that cortisol is elevated after long-term alcoholism (E, from (Stalder et al., 2010)), and that alcohol preference is reduced in adrenalectomized animals and restored by cortisol replacement (F, data from (Fahlke et al., 1994)).

## **Results**

### **A model for drug-induced regulation of $\beta$ -endorphin**

To characterize the HPA opponent process, we model the regulation of  $\beta$ -endorphin (Figure 1D).  $\beta$ -endorphin is an endogenous opioid that is secreted in response to hypothalamic CRH from corticotroph cells in the pituitary gland and from POMC neurons. Alcohol taking causes CRH secretion and hence  $\beta$ -endorphin secretion (Čupić et al., 2017; Lee et al., 2001, 2004; Rivier, 1996; Rivier et al., 1984; de Waele and Gianoulakis, 1993). In addition to causing the secretion of  $\beta$ -endorphin, CRH also activates the HPA endocrine cascade which results in the secretion of the stress hormone cortisol. Cortisol, in turn, provides negative feedback on CRH and  $\beta$ -endorphin. This feedback acts on the timescale of minutes to hours, and cannot by itself provide slow opponent properties on the timescale of days-weeks.

To address the slow timescale, we recently developed a mathematical model of the HPA axis on the time scale of weeks (Karin et al., 2020). We showed that these dynamics can explain the week-long changes in responses to CRH tests after withdrawal from prolonged alcohol use (Karin et al., 2020). Here we add  $\beta$ -endorphin to this model, and use it to study addiction.

The weeks-timescale in the model is due to growth of the total mass of the pituitary corticotroph cells that secrete ACTH and  $\beta$ -endorphin, and the adrenal cortex cells that secrete cortisol. Hereafter, we call such total cell masses “gland masses” for brevity, although the glands also contain other cell types. The mass of the gland affects the amount of hormone it secretes for a given input signal: twice the mass is assumed to result in twice the hormone secretion.

These cells constantly turn over, with a turnover time of weeks. Their major growth factors are the HPA hormones themselves: CRH causes corticotroph mass to grow (Gertz et al., 1987; Westlund et al., 1985; Carey et al., 1984; Horvath, 1988; Schteingart et al., 1986; O’Brien et al., 1992; Asa et al., 1992; Bruhn et al., 1984; Young and Akil, 1985; Gulyas et al., 1991; McNicol et al., 1988), and ACTH causes the adrenal cortex mass to grow (Dallman, 1984; Lotfi and de Mendonca, 2016; Swann, 1940; Ulrich-Lai et al., 2006).

To describe the dynamics of  $\beta$ -endorphin following alcohol intake, we added equations as described in (Methods, Figure 2A). In the model, alcohol-taking is described as an input  $u_D$  that is additive with the basal circuit input  $u_B$ , so the total input is  $u = u_B + u_D$  (Figure 2A). Persistent HPA activation by alcohol increases the hormone levels, and since these hormones act as growth factors for the glands, the gland masses grow on the timescale of weeks.

To test whether the glands can implement an opponent process for  $\beta$ -endorphin secretion, we consider a months-long input pulse that corresponds to a prolonged increase in average alcohol-taking (Figure 2BC, similar results are obtained by simulating discrete drinking episodes). Initially,  $\beta$ -endorphin and cortisol levels increase, because the drug activates the HPA axis. The increase in  $\beta$ -endorphin after drug intake is the ‘a-process’ in the opponent-process model, and has a timescale of a few hours. This occurs without significant change in the gland masses.

Over the next several weeks, glands masses adjust, as described in Karin et al. (Karin et al., 2020). A change in gland mass is assumed to cause a proportional change in the rate of secretion of the hormone produced by the gland. This is because we consider the secretion rate of each hormone as a product of three factors: the effect of the signals (agonists like CRH and antagonists like glucocorticoids), the maximal secretion capacity per unit biomass, and the total biomass of the cells. These factors are separated by timescales: signals can change over minutes, maximal secretion capacity over hours (due to protein expression changes), and total cell biomass over days-weeks (due to hypertrophy and hyperplasia). Thus, at a given level of signals and at a given maximal secretion capacity per unit biomass, the production of the hormone is proportional to the cell total biomass: doubling the biomass doubles secretion. The model takes into account both the fast changes in signals and the slow changes in gland masses. The separation into these factors allows us to consider the effects over weeks of gland mass changes and disentangle them from faster effects. It also allows us to follow changes in the mass of two cell types- pituitary corticotrophs and adrenal cortex cells - which change at the same time.

With these considerations, we see that HPA activation causes growth of pituitary corticotroph mass over weeks, due to the action of CRH as a growth factor. The enlarged biomass enhances  $\beta$ -endorphin secretion. In parallel, HPA activation makes the adrenal mass grow because of increased levels of its growth factor, ACTH. Growth of the adrenal mass causes increased secretion of glucocorticoids, which eventually suppresses  $\beta$ -endorphin secretion. This acts to renormalize  $\beta$ -endorphin levels back to baseline (mathematically, this is due to the integral feedback in Eq. 4-5 in Methods). The change in the adrenal mass is the slow

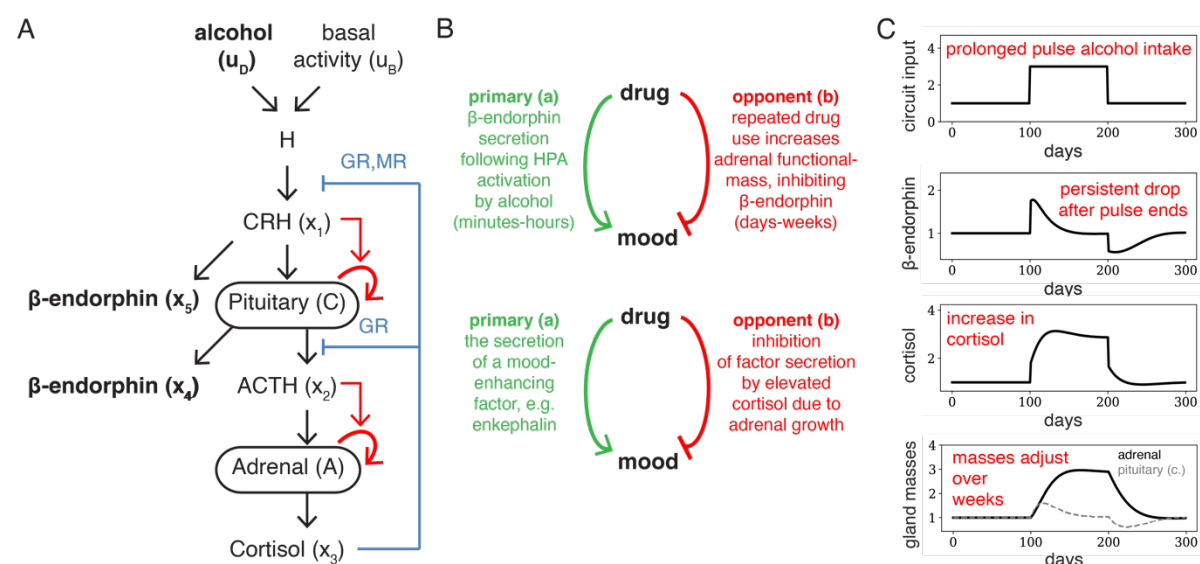


‘b-process’ in the opponent process model. It causes suppression of beta-endorphin on a timescale of days-weeks.

The same effects cause an undershoot of  $\beta$ -endorphin which lasts for weeks-months upon alcohol withdrawal after prolonged intake. The undershoot is caused by the enlarged HPA gland masses, which return to their baseline size over several weeks. These dynamics are qualitatively similar to those predicted by Gutkin et al. for the effect of nicotine on tonic dopamine through nicotinic receptors (Gutkin et al., 2006), but arise due to different physiological processes.

These opponent-process stages of gland growth and shrinkage occur regardless of model parameters, as long as the turnover time of the gland cells is much slower than the half-lives of the hormones. The opponent process is thus a robust prediction of the model.

The model explains several important observations regarding alcohol addiction. It explains the large increase in average cortisol levels seen after prolonged drinking (Stalder et al., 2010), the acute elevation of  $\beta$ -endorphin levels after alcohol administration (Frias et al., 2002; Marinelli et al., 2003), and the prolonged drop seen in  $\beta$ -endorphin levels after alcohol withdrawal (Aguirre et al., 1990; Esel et al., 2001; Inder et al., 1995; Kiefer et al., 2002; Marchesi et al., 1997; Vescovi et al., 1992). More generally, it captures the dynamics of dysregulation of HPA hormones during alcohol withdrawal (von Bardeleben et al., 1989; Karin et al., 2020). These effects are due in the model to the slow changes in the gland masses.



**Figure 2. Opponent process for  $\beta$ -endorphin based gland-mass changes in the HPA axis.** (A) We model HPA control of  $\beta$ -endorphin by adding  $\beta$ -endorphin to our recent model of the HPA axis (Karin et al., 2020) that incorporated week-timescale interactions for the gland masses (red arrows). CRH causes the growth of pituitary corticotroph cells, and ACTH causes the growth of the adrenal cortex cells. The model equations and parameters are provided in Methods. (B) The model provides opponent process properties, where the primary process is the secretion of  $\beta$ -endorphin following HPA activation, and the opponent process is changes in cortisol secretion due to change in adrenal mass



following prolonged HPA activation. The results generalize to other mood-enhancing factors that are secreted in response to alcohol intake, and that are inhibited by cortisol. (C) A prolonged increase in alcohol-taking causes  $\beta$ -endorphin to initially increase (primary process) but then return to baseline (opponent process). The reason for this is the growth of the gland masses, which increases the negative feedback of cortisol on  $\beta$ -endorphin. The return of  $\beta$ -endorphin to baseline develops over the timescale of weeks, which is much slower than the half-life of  $\beta$ -endorphin. Stopping alcohol intake after months leads to a drop in  $\beta$ -endorphin levels which lasts for a few weeks, together with an undershoot in pituitary corticotroph mass. Shown is pituitary  $\beta$ -endorphin; similar secretion patterns occur for  $\beta$ -endorphin secreted from POMC neurons.

### Opponent-process behavior is due to a fold-change detection property

The opponent-process behavior of  $\beta$ -endorphin is due to a systems-level feature of the HPA model called fold-change detection (FCD, Figure 3A). This property has been extensively studied in systems biology. A system with FCD has an output whose entire dynamics depends only on *relative* changes in input, rather than absolute changes (Adler and Alon, 2018; Shoval et al., 2010). All systems with FCD also have the property of exact adaptation, where the output adapts precisely back to baseline after step changes in its input (Ferrell, 2016; Shoval et al., 2010).

The HPA model shows that the dynamic response of CRH depends only on relative changes in input, rather than absolute changes (see Methods for the proof). Because  $\beta$ -endorphin is secreted in response to CRH, it also has the FCD property (see Methods for the proof both for pituitary and brain secretion of  $\beta$ -endorphin). The FCD property holds at low to moderate levels of HPA-axis activation, when the low-affinity GR activation is weak. At higher levels of HPA activation, GR feedback ameliorates FCD and causes reduced responses to the same fold-changes in input.

The reason for FCD is that gland masses grow in response to an increase in input on the timescale of weeks. The increased gland mass results in higher levels of cortisol. Cortisol levels thus rise proportionally to the input level. Since cortisol inhibits CRH and  $\beta$ -endorphin secretion, the latter are “normalized” by the input and show exact adaptation back to their baseline (Figure 2B), as well as fold-change responses (Figure 3A).

The FCD property provides the essential hallmarks of an opponent process (Figure 3B), as defined in (Koob and Volkow, 2010). These hallmarks concern the hedonic responses to single and repeated episodes of drug intake. After drug intake, a positive hedonic response occurs, and matches the intake’s duration and intensity. The negative hedonic response, which is due to the opponent process, follows the positive response, with a slow and prolonged decay. Repeated exposure to the drug reduces the positive response and increases the negative response, so the overall average hedonic response remains constant (this is referred to as reward allostasis).

All of these hallmarks are captured by FCD (Figure 3B). The positive response of an FCD circuit to a pulse input is always followed by a slow trailing negative response, because the overall integral of the response must be zero (Figure 3B, proof in Methods). Because FCD circuits adapt to the average input level, repeated exposure to drug intake desensitizes the positive response and sensitizes the negative response, leading to reward allostasis (Figure 3B). We therefore propose that FCD, a concept from systems biology that is common in sensory circuits, aligns with the essential features of the opponent process theory.

To make these notions more precise, we attempted to connect these FCD properties to the reward aspects of the opponent process theory more formally. For this purpose, we need to operationally define three variables in quantitative terms: subjective reward, drug reward, and withdrawal mood. We denote by  $u_B$  the basal input to the HPA axis and by  $u_D(t)$  the input due to drug intake at time  $t$ , so the total input to the HPA axis is  $u(t) = u_B + u_D(t)$ . We take the generic assumption that the effect of  $\beta$ -endorphin on subjective reward is a saturating (Hill) function, which describes the saturation of the opioid receptors, or downstream saturation of subjective reward in the brain. Thus, the instantaneous subjective reward is  $r(t, u) = \frac{1}{1+\beta^{-n}}$  where  $\beta$  is the level of  $\beta$ -endorphin and  $u$  is the HPA axis input is.

The total subjective reward  $R$  from taking a drug can be calculated by the accumulated subjective reward over a time period after taking the drug. As customary in the reinforcement learning literature, total subjective reward is given by the discounted integral over the instantaneous reward:  $R(t, u) = \int_{\tau=t}^{\infty} \gamma^{\tau} r(\tau, u) d\tau$  where  $\gamma < 1$  is a ‘future discounting’ factor. Individuals adjust drug intake over time to maximize  $R(t, u)$ . We assume that the drug reward is the total subjective reward  $R(t, u_B + u_D(t))$ , and that the mood during withdrawal from the drug is the total subjective reward without drug  $R(t, u_B)$ .

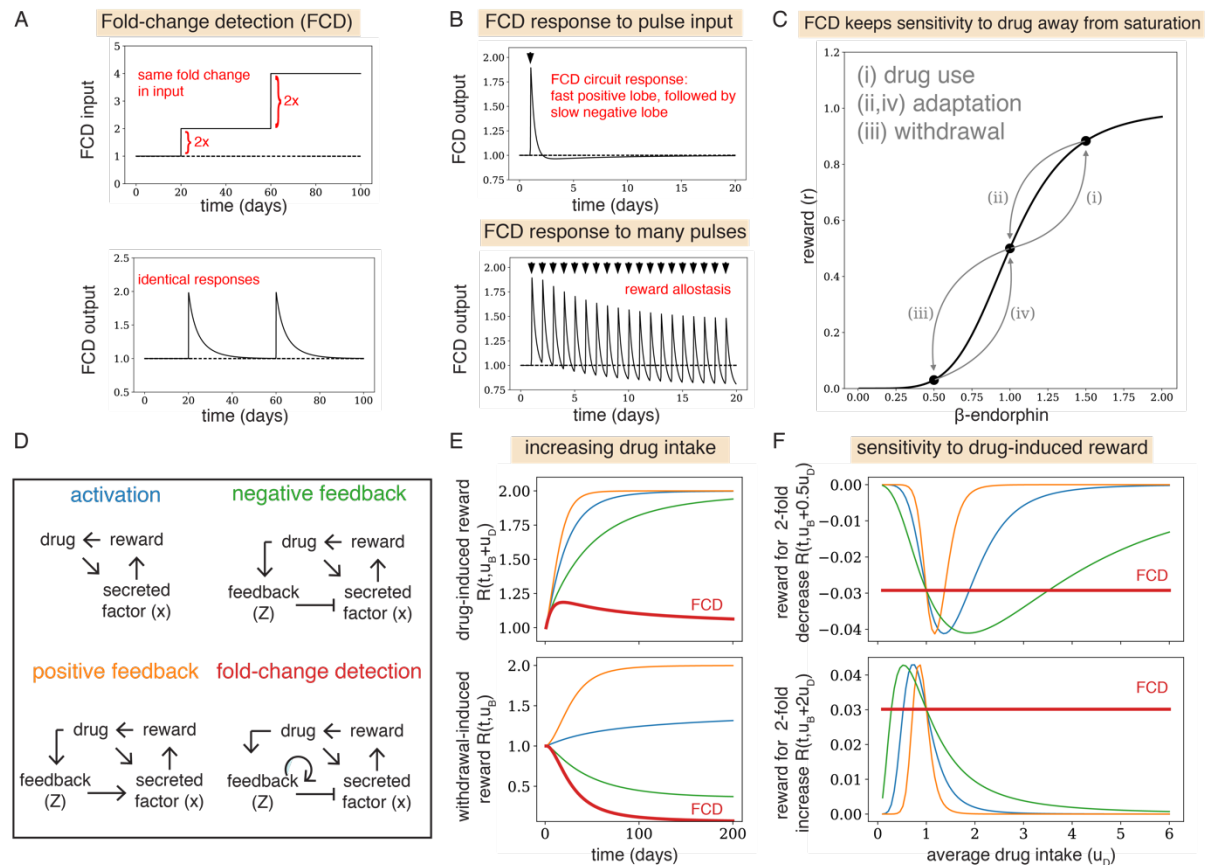
With these definitions, the changes in subjective reward in the addiction process are illustrated schematically in Figure 3C. Initially, drug use, which transiently increases  $\beta$ -endorphin levels, causes an increase in subjective reward (Figure 3C, arrow (i)). This is the initiation phase. After a few weeks, adaptation due to FCD brings  $\beta$ -endorphin and subjective reward back to baseline, despite the increase in drug intake (Figure 3C, arrow (ii)). The circuit therefore produces tolerance to the subjective reward of the drug, leading to escalation of drug intake. Withdrawal from the drug causes a drop in subjective reward (Figure 3C, arrow (iii)) which resolves after several weeks (Figure 3C, arrow (iv)). We conclude that the  $\beta$ -endorphin circuit, in which  $\beta$ -endorphin responds to fold-changes in input, shows the hallmarks of the opponent-process model of addiction.

## **FCD circuits are more vulnerable to addiction than other circuits**

FCD thus seems to provide the hallmarks for addiction, by providing a fast response and a slow opponent process. We asked whether FCD is special in this regard, or whether other common circuit designs in biology are also prone to addiction. To test this, we compare different prototypical circuits (Figure 3D) in which a drug induces the secretion of a factor which increases subjective reward. We compare the FCD circuit with other circuits that have a slow component with either negative or positive feedback on the secreted factor. We also consider a simple activation topology without a slow component. To make a ‘mathematically controlled comparison’ (Savageau, 1972), we provide all circuits with the same timescale parameters (Methods).

Simulations show that the FCD circuit is unique in the sense that drug reward becomes dissociated from drug intake level and remains constant, and that withdrawal mood drops without bound (Figure 3E). These properties can be shown analytically for all circuits (Methods). Therefore, FCD best captures the properties of the opponent-process model.

To intuitively understand why FCD is especially fragile to addiction, consider the following explanation. At any level of drug intake, the FCD circuit shows a constant (positive) preference for increasing drug intake by 2-fold, and a constant (negative) preference for decreasing drug intake by 2-fold (Figure 3F). This is in contrast with the other circuits, where preference for increasing drug intake, and dislike for decreasing drug intake, both diminish at high levels of drug intake, due to saturation effects. The reason for this is the FCD dynamics presented in Figure 3C, arrow (ii): as drug intake increases, FCD constantly pushes the sensitivity to further intake away from saturation, thus motivating further increases in intake, and demotivating decreases in intake. This may explain the fragility of FCD to addiction, since increasing drug intake does not diminish preference for further increases in drug intake.



**Figure 3. FCD control of subjective reward provides the hallmarks of the opponent process model and is prone to addiction** (A) Fold-change detection is a property of biological circuits where, after adaptation, the circuit responds to relative rather than absolute changes in input. (B) A pulse input (dark arrow) causes a doubled lobed response with a fast-positive response followed by a slow-negative response, so the overall integral is zero. This property entails that positive hedonic responses (“a-process”) are balanced by subsequent negative hedonic responses (“b-process”). Repeated pulses desensitize the positive response and sensitize the negative response so the average overall output level is constant, explaining reward allostasis. (C) FCD circuits maintain sensitivity to drug intake by keeping the system away from saturation. We demonstrate this with a schematic illustration, using a generic hill-function which maps between  $\beta$ -endorphin levels and subjective reward  $r = \frac{1}{1+\beta^{-5}}$ . Initially, the individual will increase drug intake to maximize subjective reward. This causes  $\beta$ -endorphin to increase, so that subjective reward rises closer to saturation (arrow marked (i)). Then, due to FCD,  $\beta$ -endorphin secretion adapts back to baseline over weeks, and subjective reward returns to baseline (ii). The circuit maintains its sensitivity relative to the new baseline, so that a larger drug increase is required to get the same rise in subjective reward. Finally, when the individual ceases to consume the drug,  $\beta$ -endorphin and subjective reward fall (iii), until  $\beta$ -endorphin re-adapts and returns to baseline over weeks (iv). (D) Four prototypical circuits for the control of subjective reward, FCD (red lines) where subjective reward depends on the fold-change of drug input, negative or positive feedback (green and orange lines, respectively), where a slow process Z feeds back on subjective reward, and activation (blue lines), where the drug input translates directly into subjective reward (equations in Methods). (E) Simulations of drug intake that rises linearly with time,  $u(t) = 1 + 0.015t$ . In all circuits except FCD, drug-induced reward rises with the increase in drug intake. Withdrawal mood drops without bound only for the FCD circuit. (F) The FCD circuit (red line) maintains preference for an increase in drug intake even at high levels of drug intake, whereas other circuits show diminishing preference at high intake levels.

### FCD may be advantageous for learning and reward shaping

Finally, we ask what might be the selective advantage of an FCD circuit for reward. We therefore analyze possible advantages of FCD in circuits that control subjective reward.

We focus on the role of reward in the learning of useful behaviors, a major task for cognitive systems. We demonstrate two advantages that FCD confers on learning of behaviors.

Behavior is learned by adjusting actions according to outcomes, such that actions that lead to rewarding outcomes are repeated more often (that is, they are reinforced); actions that lead to aversive outcomes are avoided (Dickinson, 1994; Herrnstein, 1970; Thorndike, 1927). This principle, known as *Thorndike's law of effect*, is a cornerstone of our understanding of animal learning. It has been mathematically analyzed in the field of reinforcement learning (Dayan and Daw, 2008; Sutton and Barto, 2018).

From the point of view of learning, FCD control of subjective reward is initially puzzling. Algorithms for reinforcement learning, such as the widely-used Q-learning algorithm for optimal action choice, often assume a direct correspondence between the value of an input and its translation into subjective reward. FCD, on the other hand, causes the subjective reward to depend also on the background level of input, breaking the direct correspondence between the input stimulus and its rewarding properties.

Here, we point out that FCD can be crucial for reinforcement learning when rewards span several orders of magnitude. By rescaling inputs according to the background input, FCD allows learning despite large differences in reward values. This FCD feature provides a wide dynamic range over decades of input. An analogous behavior over multiple scales is provided by FCD in systems such as the *E. coli* chemotaxis circuit (Adler and Alon, 2018).

The importance of such scale-invariant sensing for learning has been demonstrated for artificial algorithms such as neural networks (Ioffe and Szegedy, 2015; Santurkar et al., 2018; Sola and Sevilla, 1997). Scale invariance makes gradients more reliable and predictable, reducing the likelihood of vanishing or exploding gradients, and reduces sensitivity to hyperparameters and initialization (Santurkar et al., 2018).

In fact, artificial reinforcement learning algorithms have been shown to benefit from FCD-like processes. Van Hasselt et al. recently added adaptive normalization of the target values of a deep Q-network, which allowed the algorithm to learn to play computer games with varying magnitudes of reward (van Hasselt et al., 2016). This adaptive normalization is similar to FCD, since the rewards were effectively normalized by their background level. This raises the possibility that scale-invariance and FCD may also be important for biological reinforcement learning.

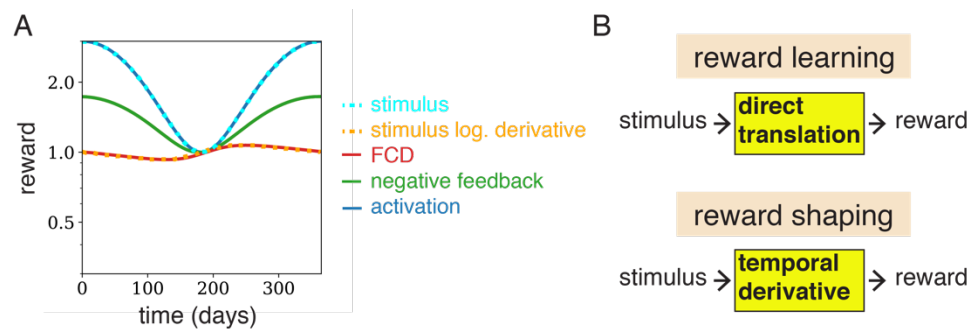
In addition to scale-invariance, FCD has a second important benefit: it provides a natural implementation of a well-established concept in reinforcement learning – *potential-based reward shaping* (Devlin and Kudenko, 2011, 2012; Laud, 2004; Ng et al., 1999; Skinner,

2019). Reward shaping is the addition of auxiliary rewards in order to guide exploration and behavior towards desirable objectives. The motivation behind reward shaping is that ‘real’ rewards, such as access to food or sexual contact, may be sparsely achieved. This sparsity limits the extent to which an agent can learn how to achieve these rewards.

As an example, consider a task presented by Ng. et al (Ng et al., 1999; Randsjøv and Alstrøm, 1998), where an agent needs to learn to ride a bicycle from point *A* to a distant point *B*. To learn this, the agent receives a reward upon reaching *B*. Since reaching *B* is difficult, the agent gets only sparse feedback for its performance. To address this, the agent is also provided with an auxiliary reward when it approaches *B* (Randsjøv and Alstrøm, 1998). However, this leads to a problem, since the agent can now receive reward by simply riding in loops around *B*, gaining the proximity reward again and again, without ever reaching *B*.

To solve this problem, and provide auxiliary rewards that preserve the learning of optimal behavior, Ng et al. suggested that the agent must be provided with *potential-based auxiliary rewards* (Ng et al., 1999). The intuition is that potential functions have a net effect of zero on any loop. Their net effect depends only on the start and end points. This concept was extended by Devlin and Kudenko (Devlin and Kudenko, 2012) to dynamical potential-based auxiliary rewards. In the standard notation of reinforcement learning, the auxiliary reward for a transition from state *s* at time *t* to state *s'* at time *t'* is given by  $F(s, t, s', t') = \gamma \Phi(s', t') - \Phi(s, t)$ , where  $\gamma$  is a discounting factor and  $\Phi$  is a real-valued potential function. Potential-based reward shaping does not have a problem of loops that can distract the agent. Moreover, potential-based reward shaping preserves the learning of optimal behavior (or *policy invariance*) (Ng et al., 1999).

FCD control of subjective reward can provide a physiological implementation of potential-based reward shaping (Figure 4). The reason for this is that the output of the FCD circuit tracks the logarithmic derivative of its input (Adler and Alon, 2018; Adler et al., 2014; Lang and Sontag, 2016). Thus FCD provides a potential function equal to the log of the input. (see Methods). Such circuits can therefore be thought of as circuits to guide exploration, rather than to learn values.



**Figure 4. FCD outputs the logarithmic derivative of the input, and therefore provides a physiological implementation of potential-based reward shaping.** (A) The negative feedback circuit output (green) and activation circuit output (blue) track the input stimulus (cyan), whereas the FCD circuit output (red) tracks its logarithmic derivative (orange). The input used is  $u(t) = e^{-\cos(\frac{2\pi t}{t_{max}})}$ ,  $t_{max} = 1\text{year}$ . The logarithmic derivative of this stimulus is  $d = (\log(u(t)))' = \frac{2\pi}{t_{max}} \sin(\frac{2\pi t}{t_{max}})$ , shown scaled and shifted for clarity,  $\frac{d+1}{w_z}$  (see methods). (B) Physiological circuits like those discussed in this study convert environmental stimuli to subjective reward. Circuits where the reward corresponds directly to the input stimuli (such as direct activation, negative, or positive feedback) facilitate learning of the input stimuli. On the other hand, circuits like FCD, where the subjective reward tracks the derivative of the input stimuli, facilitate reward shaping, where input stimuli is not directly learned but instead guides exploration towards rewarding behavior.

## Discussion

We propose an opponent process in alcohol addiction based on dynamical changes in the gland masses of the HPA axis. The gland-mass changes due to chronic alcohol intake cause the initial rise in  $\beta$ -endorphin to settle back down to baseline within weeks, and to undershoot for weeks after withdrawal. These dynamics contribute to the physiological basis of the hedonic tolerance and withdrawal stages of addiction. The present opponent process may also apply to other addictive drugs that activate the HPA axis, including cocaine and nicotine (Armario, 2010). This suggests that, in addition to neurotransmitter circuits, endocrine gland mass changes may be potentially important for addiction.

Of particular mechanistic importance in the model is the growth of the adrenal cortex caused by chronic activation of the HPA axis. This growth increases cortisol secretion, which suppresses  $\beta$ -endorphin secretion. Consistent with the model, cortisol levels are about 3-fold higher in people with alcohol abuse disorder during active drinking periods (Stalder et al., 2010); cortisol returns to normal levels after several weeks of abstinence (von Bardeleben et al., 1989; Esel et al., 2001; Marchesi et al., 1997). The model explains why adrenalectomy diminishes alcohol preference in rodents (Fahlke, 2000; Fahlke et al., 1994; Goeders, 2002; Goeders and Guerin, 1996; Hansen et al., 1995; Lamblin and De Witte, 1996). However, the role of HPA gland masses in alcohol abuse disorder or other drug addictions has not, to the best of our knowledge, been directly studied, except for Carsin-Vu et al. (Carsin-Vu et al., 2016) which provide evidence for adrenal enlargement in alcohol abuse disorder.



While our model focused on the role of  $\beta$ -endorphin, it can potentially be generalized to other factors important to alcohol addiction. Any secreted factor that is stimulated by alcohol and inhibited by glucocorticoids, or is secreted as a response to hypothalamic CRH, is predicted by the model to have similar long-term dynamics to  $\beta$ -endorphin (Methods). One potentially relevant factor is dopamine secretion, which is affected by glucocorticoids through interactions with several neural systems (Butts et al., 2011; Piazza and Le Moal, 1996; Piazza et al., 1996). Additional factors include enkephalins which preferentially bind the delta opioid receptor (Froehlich et al., 1991; Marinelli et al., 2005) and are implicated in alcohol addiction (Figure 2B). Future extensions of the present model may include these factors.

The HPA-activating effect of drugs such as alcohol provides reward for mild to moderate stimulation. However, at high levels, limiting mechanisms begin to act. One mechanism is the inhibition of CRH secretion by high levels of cortisol through the low-affinity glucocorticoid receptor GR. Another mechanism is secretion of dynorphins, endogenous opioids which are secreted as a result of prolonged CRH secretion (Nikolarakis et al., 1986). In contrast to  $\beta$ -endorphin, dynorphins have a dysphoric rather than a euphoric effect. This dysphoric effect can presumably prevent over-activation of the HPA axis from being rewarding. Dynorphin secretion was hypothesized to play a role in the dysphoria that follows drug withdrawal (Koob and Volkow, 2016).

The present model infers that increased glucocorticoids during alcohol abuse act to inhibit  $\beta$ -endorphin. This inference is based on experiments done in other contexts, using pharmacological interventions and adrenalectomy. To the best of our knowledge, this inhibition effect has not been directly tested in the context of alcohol abuse disorder. Such experiments will be an important experimental test for the model.

The key mathematical feature that underlies the proposed opponent process is that  $\beta$ -endorphin responds to relative changes in average drug intake, rather than absolute changes, a feature called fold-change detection (FCD). FCD provides the essential hallmarks of the opponent process theory of addiction.

We find that FCD control of reward is therefore especially fragile to addiction. In our model for  $\beta$ -endorphin dynamics, FCD is implemented physiologically by the growth of endocrine glands. FCD in reward pathways can also be implemented in principle by other physiological mechanisms (Adler et al., 2017). For example, adaptation and relative responses to rewards (both hallmarks of FCD) are well documented for midbrain dopamine neurons (Tobler et al., 2005). FCD can occur at the cellular level by biochemical mechanisms such as receptor methylation (Lazova et al., 2011) or by gene-regulation circuits such as incoherent

feedforward loops (Goentoro et al., 2009). Regardless of the precise implementation, FCD promotes addiction by maintaining sensitivity to increasing drug intake levels, keeping them away from reward saturation. In this way, wanting the drug is never sated until limiting mechanisms kick in. The relevant timescale for the development of drug tolerance, as well as the affective withdrawal, will depend on the timescale of adaptation in the FCD circuit. While this may be days to weeks for endocrine gland turnover, or possibly for epigenetic modifications, faster circuits (such as transcriptional circuits and protein-protein interactions) may adapt on a timescale of hours, and are therefore less likely to be important for the development of addiction on the timescale of weeks.

It will be interesting to test whether the nicotine receptor system analyzed by Gutkin et al. (Gutkin et al., 2006) also can show FCD (their original model did not generally have FCD). A recent study demonstrated how approximate FCD may be implemented in receptor-based mammalian signaling systems (Lyashenko et al., 2020). In their work, which focused on the pAkt pathway, Lyashenko et al. demonstrated that the adaptation mechanism for FCD can be implemented using receptor endocytosis. Endocytosis occurs generally in neurotransmitter pathways, and may thus be a relevant mechanism for FCD.

In addition to its vulnerability to addiction, FCD control of subjective reward also has selective benefits for learning behaviors that promote fitness. One benefit is to allow rapid learning when reward magnitude spans a large range. This is relevant when the reward baseline level is unknown or fluctuating. A second benefit is a physiological implementation of potential-based reward shaping. This is crucial when the input to the circuit represents a cue or proxy that is correlative with a ‘real’ reward, but does not have value by itself.

Not all behaviorally relevant circuits have FCD: for example, mechanical pain does not adapt to background level (and hence cannot have FCD which entails exact adaptation), whereas pain mediated by interaction with the capsaicin receptor does (Holzer, 1991; Nolano et al., 1999; Winter et al., 1995; Yao and Qin, 2009). It would be fascinating to compare the design principles of different physiological circuits that control subjective reward.

In summary, we propose an opponent process for addiction based on gland-mass changes in the HPA axis. To test this, further experiments are important. Monitoring gland masses using imaging during the stages of the addiction process can help to test this proposal. If gland masses turn out to be important for the week-scale dynamics of addiction, they might serve as relevant targets for intervention. Interventions that suppress gland mass changes at the right time, perhaps using HPA agonists and antagonists, are predicted to interfere with the addiction process and to reduce withdrawal symptoms.

## **Materials and Methods**

### **Model for $\beta$ -endorphin control by the HPA axis with gland-mass changes**

The HPA axis is a cascade of hormones activated by stress inputs. It is initiated by the secretion of CRH from the paraventricular nucleus of the hypothalamus, in response to drugs, as well as to physiological and psychological stresses, whose combined effect is the input denoted  $u$ . CRH causes pituitary corticotrophs to secrete ACTH, which, in turn, causes cortisol secretion from the adrenal cortex. CRH also acts as a growth factor for pituitary corticotrophs, and ACTH is a growth factor for the adrenal cortex. To model these dynamics, we used our recent model (Karin et al., 2020), which added to the classic HPA model of Andersen et al. (Andersen et al., 2013) two equations for the changes in gland masses. The non-dimensionalized equations are:

$$\frac{dx_1}{dt} = w_1(g_1(x_3)u - x_1) \quad [1]$$

$$\frac{dx_2}{dt} = w_2(Cg_2(x_3)x_1 - x_2) \quad [2]$$

$$\frac{dx_3}{dt} = w_3(Ax_2 - x_3) \quad [3]$$

$$\frac{dC}{dt} = w_C C(x_1 - 1) \quad [4]$$

$$\frac{dA}{dt} = w_A A(x_2 - 1) \quad [5]$$

Where  $x_1, x_2, x_3$  are the concentrations of the hormones CRH, ACTH, and cortisol,  $C$  is corticotroph functional mass, and  $A$  is the adrenal cortex functional mass. Equations 1-3 for hormone dynamics have a timescale of hours (given by the turnover rates  $w_1, w_2, w_3$ ). Equations 4-5 have a timescale of weeks due to the gland turnover rates  $w_C$  and  $w_A$ . All parameters are provided in Table 1. Finally, cortisol has negative feedback on CRH due to both the low affinity receptor GR:  $G(x_3) = \frac{1}{1 + \left(\frac{x_3}{K_{GR}}\right)^{n_{GR}}}$  (Andersen et al., 2013; Gupta et al., 2007),

and the high affinity receptor MR. Because cortisol saturates MR at physiological levels (Andersen et al., 2013), it is modelled as a Michaelis-Menten function in its saturated regime  $M(x_3) = 1/x_3$ . Cortisol also has negative feedback on the pituitary due to GR. Therefore  $g_1(x_3) = M(x_3)G(x_3)$ , and  $g_2(x_3) = G(x_3)$ .

In the present study, we added two new equations for the dynamics of  $\beta$ -endorphin. There are two  $\beta$ -endorphin sources. The first is  $\beta$ -endorphin secreted from pituitary corticotrophs into the bloodstream. This  $\beta$ -endorphin, which we denote by  $x_4$ , is co-regulated with ACTH, and therefore is described by the equation:

$$\frac{dx_4}{dt} = w_4(Cg_2(x_3)x_1 - x_4) \quad [6]$$

The second case is  $\beta$ -endorphin that is secreted into the cerebrospinal fluid from POMC neurons in the hypothalamus (Bloch et al., 1978; Bloom et al., 1978) in response to CRH, which we denote by  $x_5$ :

$$\frac{dx_5}{dt} = w_5(x_1 - x_5) \quad [7]$$

Eq. 7 can also describe  $\beta$ -endorphin secretion from non-corticotroph cells in the pituitary.

For future studies, one may propose a more general equation for a secreted factor  $x_6$  that is stimulated by alcohol and inhibited by cortisol:

$$\frac{dx_6}{dt} = w_6(g_1(x_3)u - x_6) \quad [8]$$

### FCD in $\beta$ -endorphin dynamics

We now show that  $x_4, x_5, x_6$  all have the FCD property for low to moderate levels of stress input, namely when GR is not appreciably activated,  $x_3 \ll K_{GR}$ . FCD is defined by dynamics of  $x_4, x_5$  in response to an input  $\lambda u(t)$  that are independent of  $\lambda > 0$ , starting from initial conditions at steady-state of all the hormones and glands. To show this, we use normalized variables  $\tilde{x}_3 = \lambda^{-1}x_3$  and  $\tilde{A} = \lambda^{-1}A$  in Eq. 1-8. Since  $x_3 \ll K_{GR}$ , the contribution of the GR to the negative feedback is negligible ( $G(x_3) \approx 1$ ), and thus  $g_1(x_3) \approx 1/x_3$  and  $g_2(x_3) \approx 1$ . We therefore get:

$$\frac{dx_1}{dt} = w_1\left(\frac{\lambda u}{x_3} - x_1\right) = w_1\left(\frac{u}{\tilde{x}_3} - x_1\right) \quad [9]$$

$$\frac{dx_2}{dt} = w_2(Cx_1 - x_2) \quad [10]$$

$$\frac{d\tilde{x}_3}{dt} = \lambda^{-1}\frac{dx_3}{dt} = \lambda^{-1}w_3(Ax_2 - x_3) = w_3(\tilde{A}x_2 - \tilde{x}_3) \quad [11]$$

$$\frac{dC}{dt} = w_C C(x_1 - 1) \quad [12]$$

$$\frac{d\tilde{A}}{dt} = \lambda^{-1}\frac{dA}{dt} = w_A\lambda^{-1}A(x_2 - 1) = w_A\tilde{A}(x_2 - 1) \quad [13]$$

$$\frac{dx_4}{dt} = w_4(Cx_1 - x_4) \quad [14]$$

$$\frac{dx_5}{dt} = w_5(x_1 - x_5) \quad [15]$$

$$\frac{dx_6}{dt} = w_6\left(\frac{\lambda u}{x_3} - x_6\right) = w_6\left(\frac{u}{\tilde{x}_3} - x_6\right) \quad [16]$$

Eq. 8-14 are the same as Eq. 1-7, but without the scale of the input  $\lambda$ . In addition, the initial conditions of the hormones are independent of  $\lambda$ , since at steady-state  $x_1 = x_2 = x_4 = x_5 = x_6$ , and  $x_3 = \lambda u$  so that  $\tilde{x}_3 = 1$ . The initial conditions for the gland masses are also independent of  $\lambda$ , since at steady state  $C = 1$  and  $A = \lambda u$ , so  $\tilde{A} = 1$ . Since both equations and initial conditions are independent of  $\lambda$ , the dynamics of  $x_1, x_2, x_4, x_5, x_6$  (as well as  $C$ ) are independent of the scale of the input,  $\lambda$ : we obtain the FCD property. Note that this holds only when  $x_3 \ll K_{GR}$ , otherwise the responses become blunted by the action of GR (the opponent-process property, however, still hold since the negative feedback from cortisol is further strengthened).

### Alternative circuits

Here we provide the equations for the four circuits of Figure 3. In all circuits, we model subjective reward with a generic hill equation:  $r = \frac{1}{1+x^{-n_r}}$ , where  $x$  is the factor that activates subjective reward (akin to  $\beta$ -endorphin). The hill-coefficient  $n_r$  determines the sensitivity to changes in  $x$ . For the direct-activation circuit, input  $u$  directly increases  $x$ . For the other circuits,  $x$  is activated by  $u$  and modulated by a slow variable  $Z$ . The slow variable  $Z$ , which provides feedback on the secretion of  $x$ , is analogous to the role of the gland functional masses in the HPA axis model (for simplicity of analysis and generality we consider only a single functional mass). The equations for the circuits analyzed in Fig. 3DEF are as follows.

Direct activation circuit:

$$\dot{x} = w_x(u - x) \quad [17]$$

Positive feedback circuit:

$$\dot{x} = w_x(uZ^{-1} - x) \quad [18]$$

$$\dot{Z} = w_Z(x - Z) \quad [19]$$

Negative feedback circuit:

$$\dot{x} = w_x(uZ^{0.5} - x) \quad [20]$$

$$\dot{Z} = w_Z(x - Z) \quad [21]$$

FCD circuit:

$$\dot{x} = w_x(uZ^{-1} - x) \quad [22]$$

$$\dot{x} = w_Z Z(x - 1) \quad [23]$$

All circuits:

$$r = \frac{1}{1 + x^{-n_r}} \quad [24]$$

For all circuits  $w_x = 10 \text{ min}^{-1}$ ,  $w_Z = 5 \text{ day}^{-1}$ ,  $n_r = 5$ . The precise values of the powers of  $Z$  in the equations do not affect the qualitative results.

### Simulations for FCD responses to steps, pulses, and train of pulses.

To generate the simulations in Fig. 3AB, we used the FCD circuit described above,  $\dot{x} = w_x \left( \frac{u}{Z} - x \right)$ ,  $\dot{Z} = w_Z Z(x - 1)$ , with  $w_x = 10 \text{ min}^{-1}$ ,  $w_Z = 5 \text{ day}^{-1}$ ,  $n_r = 5$ . The FCD output plotted is  $x$ , which is analogous to  $\beta$ -endorphin in the HPA axis, and  $Z$  is analogous to the adrenal mass in the HPA axis. The input in Fig. 2A consists of two step increases, from  $u=1$  to  $u=2$  at day 20, and from  $u=2$  to  $u=4$  at day 60. In Fig. 3B, the input is given by  $u(t) = u_B + u_D$ , where  $u_B = 1$  and alcohol intake is  $u_D(t) = \sum_{t_i \in T} \text{step}(t > t_i) e^{-\gamma_{\text{alcohol}}(t-t_i)}$  where  $T = \{t_0, \dots, t_{\text{max}}\}$  are the times when alcohol intake occurred. Here we took an alcohol half-life of 6 hours,  $\gamma_{\text{alcohol}} = \frac{\log 2}{6 \cdot 60} \text{ min}^{-1}$ . For the upper panel we used  $T = \{1\}$  (in units of days) and for the lower panel  $T = \{1, 2, \dots, 19\}$ . Reward allostasis is achieved because of the slow changes in  $Z$ , which cause the average of  $x$  to be kept around 1. To see this, consider an infinite pulse train and two time points  $t_1 < t_2$ , that both occur before the beginning of a pulse, where  $t_1$  is large enough so that  $Z(t_1) = Z(t_2)$ . Then:  $\langle x \rangle = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} x \, dt = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \left( \frac{\dot{Z}}{w_Z Z} + 1 \right) dt = 1$ .

### Analytical solutions for the effects of escalating drug intake on mood for various circuits

When animals are allowed to self-administrate an addictive drug such as alcohol or cocaine, they gradually escalate their intake over the timescale of days to weeks (Koob, 2013). In this section we analyze analytically how such a gradually increasing input affects the behavior of the various circuit topologies presented in Figure 3. We take a linearly increasing input  $u(t) = w_u t$  where  $w_u$  is small compared with  $w_s$  and  $w_x$ . For all circuits, when  $x$  is near its steady-state  $\frac{dx}{dt} \approx 0$ , we find  $x_{ST} = uZ^\theta$  where  $\theta = -1$  for negative feedback and FCD circuits and  $\theta = 0.5$  for the positive feedback circuit. We can substitute this expression in the equation for  $\dot{Z}$ . For FCD we get  $\dot{Z} = w_Z Z(uZ^{-1} - 1) = w_Z (u - Z)$ . Solving this equation gives:  $Z(t) = e^{-tw_Z} \left( 1 + \int_0^t e^{\tau w_Z} w_Z u(\tau) d\tau \right)$ . Substituting  $u(t) = w_u t$ , we find

$$Z(t) = e^{-tw_Z} \left( 1 + \frac{w_u}{w_Z} (1 + e^{tw_Z} (tw_Z - 1)) \right) \xrightarrow{t \rightarrow \infty} w_u t$$

And:  $x(t) = \frac{u(t)}{Z(t)} \xrightarrow{t \rightarrow \infty} 1$ . Since drug-related subjective reward  $r$  is a function of  $x$ , this suggests that even as drug use escalates, drug-induced reward drops to baseline. We next consider the effect of withdrawal, where drug use drops back to  $u = 1$ . The new steady-state of  $x$  (assuming  $Z$  is still at quasi-steady-state) is  $x_{\text{withdrawal}}(t) = \frac{u(t)}{Z(t)} \approx \frac{1}{w_u t} \xrightarrow{t \rightarrow \infty} 0$ . Withdrawal mood drops to zero in the model as drug intake escalates.

For the direct activation circuit, on the other hand,  $x$  tracks the input  $x = u$ , so drug-induced reward increases with drug intake, and withdrawal mood is the same as the original baseline. Finally, for the other circuits, we get:  $\dot{Z} = w_Z(x - Z) = w_Z Z(uZ^{\theta-1} - 1)$ , which we can solve:

$$Z(t) = \left( e^{t(-1+\theta)w_Z} + w_u \left( t + \frac{1}{(-1+\theta)w_Z} \right) \right)^{\frac{1}{1-\theta}} \xrightarrow{t \rightarrow \infty} (w_u t)^{\frac{1}{1-\theta}}$$

And:  $x(t) = uZ^{\theta} \xrightarrow{t \rightarrow \infty} (w_u t)^{1+\frac{\theta}{1-\theta}}$ , with the withdrawal level of  $x$  being  $x_{\text{withdrawal}}(t) = (w_u t)^{\frac{\theta}{1-\theta}}$ . For the negative feedback circuit, where  $\theta < 0$ , drug-induced reward increases sublinearly with the increase in time / drug intake, and withdrawal mood decreases sublinearly with drug intake / time (as the square root for  $\theta = -1$ ). Therefore, drug-induced reward remains associated with drug intake, and withdrawal mood, although decreasing, is less sensitive than the FCD circuit to the escalation in drug intake. For the positive feedback circuit, where  $0 < \theta < 1$ , drug-induced reward increases super-linearly with drug intake (quadratically for  $\theta = 0.5$ ), and withdrawal mood also increases with drug intake, in contrast with the observed phenomenology of the addiction process.

### Integral of FCD response

We consider the overall integral of the difference between subjective reward  $r$  and its baseline, to a transient perturbation given by  $u(t)$ , such as the pulse of drug intake shown in Fig. 3A. For simplicity, we assume that the changes in  $u(t)$  are small enough so that one can linearize  $r(t)$  around  $x=1$ , obtaining  $r \approx \frac{1}{2} + \frac{1}{4}n_s(x - 1)$ . The integral of the response of the FCD circuit is then:

$$\int_0^{\infty} \left( r - \frac{1}{2} \right) dt \approx \frac{n_s}{4} \int_0^{\infty} (x - 1) dt \sim \int_0^{\infty} \frac{\dot{Z}}{Z} dt = \log \left( \frac{Z(\infty)}{Z(0)} \right) = 0$$

since  $u(t)$  changes transiently,  $Z$  is the same at  $t = 0$  and at  $t = \infty$ . This is not true for the other circuits, and they may therefore have a non-zero integral response to a transient perturbation.



## Model for Reward

Reward optimization was modelled by the standard assumption that individuals try to optimize total discounted subjective reward  $R(t, u) = \int_{\tau=t}^{\infty} \gamma^{\tau} r(\tau, u)$  (Dayan and Daw, 2008; Sutton and Barto, 2018). We set  $\gamma = 0.99$ , where time is in units of minutes. This value is observed experimentally (McClure et al., 2007), however, the precise value is not important for the conclusions, as long as discounting occurs faster than the timescale of the slow component of the circuit.

## Model for reward shaping

Reward-shaping is defined within the framework of reinforcement learning, where an agent learns to predict the value of its actions (Sutton and Barto, 2018). The decision making of the agent is studied in the context of a Markov Decision Process (MDP). An MDP is a tuple  $M = \langle S, A, T, R \rangle$  where  $S$  is the state space,  $A$  is the action space,  $T(s, a, s')$  maps the probability that an action  $a$  taken at a given state  $s$  causes a transition to a different state  $s'$ , and  $R(s, a, s')$  is the (immediate) reward received for this transition. The goal of optimal learning algorithms such as Q-learning is to find a policy (a map from states to action,  $\Pi: S \rightarrow A$ ) that maximizes accumulated reward  $\sum_{t=0}^{\infty} \gamma^t r(t)$ , where  $0 \leq \gamma < 1$  is a future-discounting factor and  $r(t) = R(s_t, a_t, s_{t+1})$  is the immediate reward received at time  $t$ . This can be generalized to the continuous case, as in (Doya, 2000), by setting the accumulated reward as  $\int_{t=0}^{\infty} \gamma^t r(t) dt$ .

In order to guide learning, an additional dynamical auxiliary reward function  $F(s, t, s', t')$  can be appended, creating a new MDP  $M' = \langle S, A, T, R' \rangle$  where:

$$R'(s, a, s') = R(s, a, s) + F(s, t, s', t')$$

$F(s, t, s', t')$  is the *reward shaping* function that outputs the reward for the transition from state  $s$  at time  $t$  to state  $s'$  at time  $t'$ . The MDP  $M'$  is guaranteed to preserve the optimal policy of  $M$  if  $F(s, t, s', t')$  can be rewritten as a potential function (Devlin and Kudenko, 2012; Ng et al., 1999):

$$F(s, t, s', t') = \gamma \Phi(s', t') - \phi(s, t)$$

This because the accumulated auxiliary rewards for any sequence of states does not depend on the actions taken:

$$\begin{aligned} \sum_{t=0}^{\infty} \gamma^t (r(t) + \gamma \Phi(s_{t+1}, t+1) - \phi(s_t, t)) &= \sum_{t=0}^{\infty} \gamma^t r(t) + \sum_{t=1}^{\infty} \gamma^t \Phi(s_t, t) - \sum_{t=0}^{\infty} \gamma^t \Phi(s_t, t) \\ &= \sum_{t=0}^{\infty} \gamma^t r(t) - \Phi(s_0, 0) \end{aligned}$$

When  $\gamma \approx 1$ , the above holds approximately also when taking:

$$F(s, t, s', t') = \Phi(s', t') - \phi(s, t)$$

In the continuous case, this is equivalent to setting  $F$  as the time derivative of a function  $\Phi(s, t)$ :

$$F(s, t) = \frac{d\Phi(s, t)}{dt}$$

Let us now consider a biological circuit that translates a stimulus that depends on the current state  $u(s)$  to a subjective reward  $r$ . For circuits like direct activation, negative feedback, or positive feedback,  $r$  tracks the level of  $u(s)$ , as shown in Figure 4. The individual will therefore learn  $u(s)$  through the rewards  $r$ . On the other hand, for the FCD circuit,  $r$  tracks the *logarithmic temporal derivative* of  $u(s)$ :  $r \approx \frac{d \log(u(s))}{dt}$ . Thus,  $r$  corresponds to the temporal derivative of a potential function  $\Phi(s, t) = \log(u(s))$ . In this case, the individual will not learn  $u(s)$ , but instead  $u(s)$  is used as ‘shaping rewards’ to guide exploration through  $r$ . FCD can therefore provide a physiological implementation of reward shaping.

#### Remark on connection to social attachment

The present model may also relate to aspects of the dynamics of social relations.  $\beta$ -endorphin is thought to play a crucial role in mediating social attachment in primates (Machin and Dunbar, 2011). Social attachment has been suggested to follow distinct stages that are similar in some ways to those of drug addiction, as reviewed by Machin and Dunbar (Machin and Dunbar, 2011). An initial stage of euphoria and attraction develops into tolerance (transition from attraction to attachment). A period of sustained negative affect occurs after separation or social withdrawal. The present model may explain, at least in part, the timescales of these stages since social interactions activate the HPA axis, and therefore  $\beta$ -endorphin dynamics.

The model may also be used to study the combined effects of drug consumption and social interactions. Since both social interactions and alcohol activate the HPA axis, the model predicts that a given serving of alcohol  $u_{ALC}$  will be more rewarding for individuals with a lower baseline of social stimulus  $u_{SOC}$ , since the response is proportional to  $\frac{u_{ALC} + u_{SOC}}{u_{SOC}} = 1 + \frac{u_{ALC}}{u_{SOC}}$ . On the other hand, alcohol consumption can also raise  $\beta$ -endorphin levels during social withdrawal, which represents a drop in  $u_{SOC}$ . These observations provide testable predictions that may illuminate social aspects of drug addiction (Heilig et al., 2016).

Changes in HPA gland masses are also well documented in another disorder that is associated with chronic HPA axis activation – major depression. Depressed individuals have increased adrenal mass (Amsterdam et al., 1987; Dorovini-Zis and Zis, 1987; Dumser et al., 1998; Ludescher et al., 2008; Nemeroff et al., 1992; Rubin et al., 1996; Szigethy et al., 1994)

which returns to its original size after remission (Rubin et al., 1995). Taken together with the biological and psychological similarities between drug withdrawal and depression (Barr et al., 2002), the present study suggests that gland mass abnormalities may be a common mechanism for affective disorders. If this is indeed the case, interventions aimed at adjusting gland mass size may be beneficial for treating drug addiction and depression.

**Table 1. Parameter values for HPA model.**

Parameter	Value
$w_1$	0.17/min (Andersen et al., 2013)
$w_2$	0.035/min (Andersen et al., 2013)
$w_3$	0.009/min (Andersen et al., 2013)
$w_4$	0.0019/min (Foley et al., 1979)
$w_C$	0.099/day
$w_A$	0.049/day
$K_{GR}$	8
$n_{GR}$	3 (Andersen et al., 2013)

## References

- Adams, M.A., and Hirst, M. (1984). Adrenal and urinary catecholamines during and after severe ethanol intoxication in rats: A profile of changes. *Pharmacol. Biochem. Behav.* *21*, 125–131.
- Adler, M., and Alon, U. (2018). Fold-change detection in biological systems. *Curr. Opin. Syst. Biol.* *8*, 81–89.
- Adler, M., Mayo, A., and Alon, U. (2014). Logarithmic and Power Law Input-Output Relations in Sensory Systems with Fold-Change Detection. *PLoS Comput. Biol.* *10*, e1003781.
- Adler, M., Szekely, P., Mayo, A., and Alon, U. (2017). Optimal Regulatory Circuit Topologies for Fold-Change Detection. *Cell Syst.* *4*, 171-181.e8.
- Aguirre, J.C., Del Arbol, J.L., Raya, J., Ruiz-Requena, M.E., and Rico Irles, J. (1990). Plasma  $\beta$ -endorphin levels in chronic alcoholics. *Alcohol* *7*, 409–412.
- Amsterdam, J.D., Marinelli, D.L., Arger, P., and Winokur, A. (1987). Assessment of adrenal gland volume by computed tomography in depressed patients and healthy volunteers: a pilot study. *Psychiatry Res.* *21*, 189–197.
- Andersen, M., Vinther, F., and Ottesen, J.T. (2013). Mathematical modeling of the hypothalamic-pituitary-adrenal gland (HPA) axis, including hippocampal mechanisms. *Math. Biosci.* *246*, 122–138.
- Armario, A. (2010). Activation of the hypothalamic–pituitary–adrenal axis by addictive drugs: different pathways, common outcome. *Trends Pharmacol. Sci.* *31*, 318–325.
- Asa, S.L., Kovacs, K., Hammer, G.D., Liu, B., Roos, B.A., and Low, M.J. (1992). Pituitary corticotroph hyperplasia in rats implanted with a medullary thyroid carcinoma cell line transfected with a corticotropin-releasing hormone complementary deoxyribonucleic acid expression vector. *Endocrinology* *131*, 715–720.
- von Bardeleben, U., Heuser, I., and Holsboer, F. (1989). Human CRH stimulation response during acute withdrawal and after medium-term abstention from alcohol abuse. *Psychoneuroendocrinology* *14*, 441–449.
- Barr, A.M., Markou, A., and Phillips, A.G. (2002). A ‘crash’ course on psychostimulant withdrawal as a model of depression. *Trends Pharmacol. Sci.* *23*, 475–482.
- Berridge, K.C., and Robinson, T.E. (1998). What is the role of dopamine in reward: hedonic impact, reward learning, or incentive salience? *Brain Res. Rev.* *28*, 309–369.
- Berridge, K.C., and Robinson, T.E. (2016). Liking, Wanting and the Incentive-Sensitization Theory of Addiction. *Am. Psychol.* *71*, 670–679.
- Bloch, B., Bugnon, C., Fellman, D., and Lenys, D. (1978). Immunocytochemical evidence that the same neurons in the human infundibular nucleus are stained with anti-endorphins and antisera of other related peptides. *Neurosci. Lett.* *10*, 147–152.

- Bloom, F., Battenberg, E., Rossier, J., Ling, N., and Guillemin, R. (1978). Neurons containing beta-endorphin in rat brain exist separately from those containing enkephalin: immunocytochemical studies. *Proc. Natl. Acad. Sci.* 75, 1591–1595.
- Bruhn, T.O., Sutton, R.E., Rivier, C.L., and Vale, W.W. (1984). Corticotropin-Releasing Factor Regulates Proopiomelanocortin Messenger Ribonucleic Acid Levels in vivo. *Neuroendocrinology* 39, 170–175.
- Carey, R.M., Varma, S.K., Drake Jr, C.R., Thorner, M.O., Kovacs, K., Rivier, J., and Vale, W. (1984). Ectopic secretion of corticotropin-releasing factor as a cause of Cushing's syndrome: a clinical, morphologic, and biochemical study. *N. Engl. J. Med.* 311, 13–20.
- Carsin-Vu, A., Oubaya, N., Mulé, S., Janvier, A., Delemer, B., Soyer, P., and Hoeffel, C. (2016). MDCT Linear and Volumetric Analysis of Adrenal Glands: Normative Data and Multiparametric Assessment. *Eur. Radiol.* 26, 2494–2501.
- Čupić, Ž., Stanojević, A., Marković, V.M., Kolar-Anić, L., Terenius, L., and Vukojević, V. (2017). The HPA axis and ethanol: a synthesis of mathematical modelling and experimental observations. *Addict. Biol.* 22, 1486–1500.
- Dallman, M.F. (1984). Control of Adrenocortical Growth in Vivo. *Endocr. Res.* 10, 213–242.
- Darcq, E., and Kieffer, B.L. (2018). Opioid receptors: drivers to addiction? *Nat. Rev. Neurosci.* 19, 499–514.
- Dayan, P., and Daw, N.D. (2008). Decision theory, reinforcement learning, and the brain. *Cogn. Affect. Behav. Neurosci.* 8, 429–453.
- Devlin, S., and Kudenko, D. (2011). Theoretical considerations of potential-based reward shaping for multi-agent systems. In *The 10th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*, (International Foundation for Autonomous Agents and Multiagent Systems), pp. 225–232.
- Devlin, S.M., and Kudenko, D. (2012). Dynamic potential-based reward shaping. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems*, (IFAAMAS), pp. 433–440.
- Dickinson, A. (1994). Instrumental conditioning. In *Animal Learning and Cognition*, (San Diego, CA, US: Academic Press), pp. 45–79.
- Đikić, D., Budeč, M., Vranješ-Đurić, S., Koko, V., Vignjević, S., and Mitrović, O. (2011). The Acute Effect of Ethanol on Adrenal Cortex in Female Rats—Possible Role of Nitric Oxide. *Alcohol Alcohol* 46, 523–528.
- Doppman, J.L., Miller, D.L., Dwyer, A.J., Loughlin, T., Nieman, L., Cutler, G.B., Chrousos, G.P., Oldfield, E., and Loriaux, D.L. (1988). Macronodular adrenal hyperplasia in Cushing disease. *Radiology* 166, 347–352.
- Dorovini-Zis, K., and Zis, A.P. (1987). Increased adrenal weight in victims of violent suicide. *Am. J. Psychiatry*.

- Doya, K. (2000). Reinforcement learning in continuous time and space. *Neural Comput.* *12*, 219–245.
- Drews, E., and Zimmer, A. (2010). Modulation of alcohol and nicotine responses through the endogenous opioid system. *Prog. Neurobiol.* *90*, 1–15.
- Dumser, T., Barocka, A., and Schubert, E. (1998). Weight of adrenal glands may be increased in persons who commit suicide. *Am. J. Forensic Med. Pathol.* *19*, 72–76.
- Esel, E., Sofuoglu, S., Aslan, S.S., Kula, M., Yabanoglu, I., and Turan, M.T. (2001). Plasma levels of beta-endorphin, adrenocorticotrophic hormone and cortisol during early and late alcohol withdrawal. *Alcohol Alcohol* *36*, 572–576.
- Fahlke, C. (2000). Effect of Adrenalectomy and Exposure to Corticosterone on Alcohol Intake in Alcohol-Preferring and Alcohol-Avoiding Rat Lines. *Alcohol Alcohol* *35*, 139–144.
- Fahlke, C., Engel, J.A., Eriksson, C.J., Hård, E., and Söderpalm, B. (1994). Involvement of corticosterone in the modulation of ethanol consumption in the rat. *Alcohol Fayettev. N* *11*, 195–202.
- Ferrell, J.E. (2016). Perfect and Near-Perfect Adaptation in Cell Signaling. *Cell Syst.* *2*, 62–67.
- Foley, K.M., Kourides, I.A., Inturrisi, C.E., Kaiko, R.F., Zaroulis, C.G., Posner, J.B., Houde, R.W., and Li, C.H. (1979). -Endorphin: Analgesic and hormonal effects in humans. *Proc. Natl. Acad. Sci.* *76*, 5377–5381.
- Frias, J., Torres, J.M., Miranda, M.T., Ruiz, E., and Ortega, E. (2002). EFFECTS OF ACUTE ALCOHOL INTOXICATION ON PITUITARY–GONADAL AXIS HORMONES, PITUITARY–ADRENAL AXIS HORMONES,  $\beta$ -ENDORPHIN AND PROLACTIN IN HUMAN ADULTS OF BOTH SEXES. *Alcohol Alcohol* *37*, 169–173.
- Gerrits, M.A., Lesscher, H.B., and van Ree, J.M. (2003). Drug dependence and the endogenous opioid system. *Eur. Neuropsychopharmacol.* *13*, 424–434.
- Gertz, B.J., Contreras, L.N., McComb, D.J., Kovacs, K., Tyrrell, J.B., and Dallman, M.F. (1987). Chronic administration of corticotropin-releasing factor increases pituitary corticotroph number. *Endocrinology* *120*, 381–388.
- Gianoulakis, C. (2004). Endogenous opioids and addiction to alcohol and other drugs of abuse. *Curr. Top. Med. Chem.* *4*, 39–50.
- Glimcher, P.W. (2011). Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis. *Proc. Natl. Acad. Sci.* *108*, 15647–15654.
- Goeders, N.E. (2002). Stress and cocaine addiction. *J. Pharmacol. Exp. Ther.* *301*, 785–789.
- Goeders, N.E., and Guerin, G.F. (1996). Effects of surgical and pharmacological adrenalectomy on the initiation and maintenance of intravenous cocaine self-administration in rats. *Brain Res.* *722*, 145–152.



Goentoro, L., Shoval, O., Kirschner, M.W., and Alon, U. (2009). The incoherent feedforward loop can provide fold-change detection in gene regulation. *Mol. Cell* 36, 894–899.

Guillemin, R., Vargo, T., Rossier, J., Minick, S., Ling, N., Rivier, C., Vale, W., and Bloom, F. (1977). beta-Endorphin and adrenocorticotropin are selected concomitantly by the pituitary gland. *Science* 197, 1367–1369.

Gulyas, M., Puszta, L., Rappay, G., and Makara, G.B. (1991). Pituitary corticotrophs proliferate temporarily after adrenalectomy. *Histochemistry* 96, 185–189.

Gupta, S., Aslakson, E., Gurbaxani, B.M., and Vernon, S.D. (2007). Inclusion of the glucocorticoid receptor in a hypothalamic pituitary adrenal axis model reveals bistability. *Theor. Biol. Med. Model.* 4, 8.

Gutkin, B.S., Dehaene, S., and Changeux, J.-P. (2006). A neurocomputational hypothesis for nicotine addiction. *Proc. Natl. Acad. Sci.* 103, 1106–1111.

Hall, F.S., Sora, I., and Uhl, G.R. (2001). Ethanol consumption and reward are decreased in mu-opiate receptor knockout mice. *Psychopharmacology (Berl.)* 154, 43–49.

Hansen, S., Fahlke, C., Söderpalm, A.H., and Hård, E. (1995). Significance of adrenal corticosteroid secretion for the food restriction-induced enhancement of alcohol drinking in the rat. *Psychopharmacology (Berl.)* 121, 213–221.

Hargreaves, K.M., Schmidt, E.A., Mueller, G.P., and Dionne, R.A. (1987). Dexamethasone alters plasma levels of beta-endorphin and postoperative pain. *Clin. Pharmacol. Ther.* 42, 601–607.

van Hasselt, H.P., Guez, A., Hessel, M., Mnih, V., and Silver, D. (2016). Learning values across many orders of magnitude. In *Advances in Neural Information Processing Systems*, pp. 4287–4295.

Heilig, M., Epstein, D.H., Nader, M.A., and Shaham, Y. (2016). Time to connect: bringing social context into addiction neuroscience. *Nat. Rev. Neurosci.* 17, 592–599.

Herrnstein, R.J. (1970). On the law of effect 1. *J. Exp. Anal. Behav.* 13, 243–266.

Holzer, P. (1991). Capsaicin: cellular targets, mechanisms of action, and selectivity for thin sensory neurons. *Pharmacol Rev* 43, 143–201.

Horvath, E. (1988). Pituitary hyperplasia. *Pathol. - Res. Pract.* 183, 623–625.

Hyman, S.E., Malenka, R.C., and Nestler, E.J. (2006). NEURAL MECHANISMS OF ADDICTION: The Role of Reward-Related Learning and Memory. *Annu. Rev. Neurosci.* 29, 565–598.

Inder, W.J., Joyce, P.R., Ellis, M.J., Evans, M.J., Livesey, J.H., and Donald, R.A. (1995). The effects of alcoholism on the hypothalamic-pituitary-adrenal axis: interaction with endogenous opioid peptides. *Clin. Endocrinol. (Oxf.)* 43, 283–290.

Ioffe, S., and Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *ArXiv Prepr. ArXiv150203167*.

- Karin, O., Raz, M., Tendler, A., Bar, A., Korem Kohanim, Y., Milo, T., and Alon, U. (2020). A new model for the HPA axis explains dysregulation of stress hormones on the timescale of weeks. *Mol. Syst. Biol.* *16*, e9510.
- Kelley, A.E., and Berridge, K.C. (2002). The Neuroscience of Natural Rewards: Relevance to Addictive Drugs. *J. Neurosci.* *22*, 3306–3311.
- Kiefer, F., Horntrich, M., Jahn, H., and Wiedemann, K. (2002). Is withdrawal-induced anxiety in alcoholism based on  $\beta$ -endorphin deficiency? *Psychopharmacology (Berl.)* *162*, 433–437.
- Koob, G.F. (2011). Theoretical Frameworks and Mechanistic Aspects of Alcohol Addiction: Alcohol Addiction as a Reward Deficit Disorder. In *Behavioral Neurobiology of Alcohol Addiction*, W.H. Sommer, and R. Spanagel, eds. (Berlin, Heidelberg: Springer Berlin Heidelberg), pp. 3–30.
- Koob, G.F. (2013). Addiction is a Reward Deficit and Stress Surfeit Disorder. *Front. Psychiatry* *4*.
- Koob, G.F., and Le Moal, M. (1997). Drug abuse: hedonic homeostatic dysregulation. *Science* *278*, 52–58.
- Koob, G.F., and Le Moal, M. (2001). Drug addiction, dysregulation of reward, and allostasis. *Neuropsychopharmacology* *24*, 97–129.
- Koob, G.F., and Volkow, N.D. (2010). Neurocircuitry of Addiction. *Neuropsychopharmacology* *35*, 217–238.
- Koob, G.F., and Volkow, N.D. (2016). Neurobiology of addiction: a neurocircuitry analysis. *Lancet Psychiatry* *3*, 760–773.
- Kuzmin, A.V., Gerrits, M.A., van Ree, J.M., and Zwartau, E.E. (1997). Naloxone inhibits the reinforcing and motivational aspects of cocaine addiction in mice. *Life Sci.* *60*, PL257–PL264.
- Lamblin, F., and De Witte, P. (1996). Adrenalectomy prevents the development of alcohol preference in male rats. *Alcohol Fayettev. N* *13*, 233–238.
- Lang, M., and Sontag, E. (2016). Scale-invariant systems realize nonlinear differential operators. In *2016 American Control Conference (ACC)*, pp. 6676–6682.
- Laud, A.D. (2004). Theory and application of reward shaping in reinforcement learning.
- Lazova, M.D., Ahmed, T., Bellomo, D., Stocker, R., and Shimizu, T.S. (2011). Response rescaling in bacterial chemotaxis. *Proc. Natl. Acad. Sci.* *108*, 13870–13875.
- Lee, S., Smith, G.W., Vale, W., Lee, K.-F., and Rivier, C. (2001). Mice that lack corticotropin-releasing factor (CRF) receptors type 1 show a blunted ACTH response to acute alcohol despite up-regulated constitutive hypothalamic CRF gene expression. *Alcohol. Clin. Exp. Res.* *25*, 427–433.

- Lee, S., Selvage, D., Hansen, K., and Rivier, C. (2004). Site of action of acute alcohol administration in stimulating the rat hypothalamic-pituitary-adrenal axis: comparison between the effect of systemic and intracerebroventricular injection of this drug on pituitary and hypothalamic responses. *Endocrinology* 145, 4470–4479.
- Lotfi, C.F.P., and de Mendonca, P.O. (2016). Comparative effect of ACTH and related peptides on proliferation and growth of rat adrenal gland. *Front. Endocrinol.* 7, 39.
- Ludescher, B., Najib, A., Baar, S., MacHann, J., Schick, F., Buchkremer, G., Claussen, C.D., and Eschweiler, G.W. (2008). Increase of Visceral Fat and Adrenal Gland Volume in Women with Depression: Preliminary Results of a Morphometric Mri Study. *Int. J. Psychiatry Med.* 38, 229–240.
- Lyashenko, E., Niepel, M., Dixit, P.D., Lim, S.K., Sorger, P.K., and Vitkup, D. (2020). Receptor-based mechanism of relative sensing and cell memory in mammalian signaling networks. *ELife* 9, e50342.
- Machin, A.J., and Dunbar, R.I.M. (2011). The brain opioid theory of social attachment: a review of the evidence. *Behaviour* 148, 985–1025.
- Marchesi, C., Chiodera, P., Ampollini, P., and Volpi, R. (1997). Beta-endorphin, adrenocorticotrophic hormone and cortisol secretion in abstinent alcoholics. *Psychiatry Res.* 72, 187–194.
- Marinelli, P.W., Quirion, R., and Gianoulakis, C. (2003). A microdialysis profile of  $\beta$ -endorphin and catecholamines in the rat nucleus accumbens following alcohol administration. *Psychopharmacology (Berl.)* 169, 60–67.
- McClure, S.M., Ericson, K.M., Laibson, D.I., Loewenstein, G., and Cohen, J.D. (2007). Time Discounting for Primary Rewards. *J. Neurosci.* 27, 5796–5804.
- McKay, L.I., and Cidlowski, J.A. (2003). Pharmacokinetics of Corticosteroids. *Holl.-Frei Cancer Med.* 6th Ed.
- McNicol, A.M., Kubba, M.A.G., and McTeague, E. (1988). The mitogenic effects of corticotrophin-releasing factor on the anterior pituitary gland of the rat. *J. Endocrinol.* 118, 237-NP.
- Mendelson, J.H., Ogata, M., and Mello, N.K. (1971). Adrenal Function and Alcoholism: I. Serum Cortisol. *Psychosom. Med.* 33, 145–158.
- Milton, A.L., and Everitt, B.J. (2012). The persistence of maladaptive memory: Addiction, drug memories and anti-relapse treatments. *Neurosci. Biobehav. Rev.* 36, 1119–1139.
- Mitchell, J.M., O’Neil, J.P., Janabi, M., Marks, S.M., Jagust, W.J., and Fields, H.L. (2012). Alcohol consumption induces endogenous opioid release in the human orbitofrontal cortex and nucleus accumbens. *Sci. Transl. Med.* 4, 116ra6.
- Nemeroff, C.B., Krishnan, K.R.R., Reed, D., Leder, R., Beam, C., and Dunnick, N.R. (1992). Adrenal gland enlargement in major depression: a computed tomographic study. *Arch. Gen. Psychiatry* 49, 384–387.

- Ng, A.Y., Harada, D., and Russell, S. (1999). Policy invariance under reward transformations: Theory and application to reward shaping. In ICML, pp. 278–287.
- Nikolarakis, K.E., Almeida, O.F.X., and Herz, A. (1986). Stimulation of hypothalamic  $\beta$ -endorphin and dynorphin release by corticotropin-releasing factor (in vitro). *Brain Res.* 399, 152–155.
- Nolano, M., Simone, D.A., Wendelschafer-Crabb, G., Johnson, T., Hazen, E., and Kennedy, W.R. (1999). Topical capsaicin in humans: parallel loss of epidermal nerve fibers and pain sensation. *Pain* 81, 135–145.
- O'Brien, T., Young Jr, W.F., Davlla, D.G., Schelthauer, B.W., Kovacs, K., Horvath, E., Vale, W., and van Heerden, J.A. (1992). Cushing's syndrome associated with ectopic production of corticotrophin-releasing hormone, corticotrophin and vasopressin by a pheochromocytoma. *Clin. Endocrinol. (Oxf.)* 37, 460–467.
- Racz, I., Schürmann, B., Karpushova, A., Reuter, M., Cichon, S., Montag, C., Fürst, R., Schütz, C., Franke, P.E., Strohmaier, J., et al. (2008). The opioid peptides enkephalin and  $\beta$ -endorphin in alcohol dependence. *Biol. Psychiatry* 64, 989–997.
- Randløv, J., and Alstrøm, P. (1998). Learning to Drive a Bicycle Using Reinforcement Learning and Shaping. In ICML, (Citeseer), pp. 463–471.
- Rivier, C. (1996). Alcohol stimulates ACTH secretion in the rat: mechanisms of action and interactions with other stimuli. *Alcohol. Clin. Exp. Res.* 20, 240–254.
- Rivier, C., Brownstein, M., Spiess, J., Rivier, J., and Vale, W. (1982). In vivo corticotropin-releasing factor-induced secretion of adrenocorticotropin, beta-endorphin, and corticosterone. *Endocrinology* 110, 272–278.
- Rivier, C., Bruhn, T., and Vale, W. (1984). Effect of ethanol on the hypothalamic-pituitary-adrenal axis in the rat: role of corticotropin-releasing factor (CRF). *J. Pharmacol. Exp. Ther.* 229, 127–131.
- Robbins, T. (2002). Limbic-Striatal Memory Systems and Drug Addiction. *Neurobiol. Learn. Mem.* 78, 625–636.
- Roberts, A.J., McDonald, J.S., Heyser, C.J., Kieffer, B.L., Matthes, H.W.D., Koob, G.F., and Gold, L.H. (2000).  $\mu$ -Opioid Receptor Knockout Mice Do Not Self-Administer Alcohol. *J. Pharmacol. Exp. Ther.* 293, 1002–1008.
- Roth-Deri, I., Green-Sadan, T., and Yadid, G. (2008).  $\beta$ -Endorphin and drug-induced reward and reinforcement. *Prog. Neurobiol.* 86, 1–21.
- Rubin, R.T., Phillips, J.J., Sadow, T.F., and McCracken, J.T. (1995). Adrenal gland volume in major depression. Increase during the depressive episode and decrease with successful treatment. *Arch. Gen. Psychiatry* 52, 213–218.
- Rubin, R.T., Phillips, J.J., McCracken, J.T., and Sadow, T.F. (1996). Adrenal gland volume in major depression: relationship to basal and stimulated pituitary-adrenal cortical axis function. *Biol. Psychiatry* 40, 89–97.

Santurkar, S., Tsipras, D., Ilyas, A., and Madry, A. (2018). How does batch normalization help optimization? In *Advances in Neural Information Processing Systems*, pp. 2483–2493.

Savageau, M.A. (1972). *The Behavior of Intact Biochemical Control Systems*\* \*This will not be an exhaustive review of the different methods for analyzing biochemical systems, but rather a selective treatment of one particular approach. Reviews covering alternative approaches to these problems have recently been presented (28, 33). In *Current Topics in Cellular Regulation*, (Elsevier), pp. 63–130.

Schteingart, D.E., Lloyd, R.V., Akil, H., Chandler, W.F., Ibarra-Perez, G., Rosen, S.G., and Ogletree, R. (1986). Cushing's Syndrome Secondary to Ectopic Corticotropin-Releasing Hormone-Adrenocorticotropin Secretion\*. *J. Clin. Endocrinol. Metab.* *63*, 770–775.

Schultz, W. (1998). Predictive reward signal of dopamine neurons. *J. Neurophysiol.* *80*, 1–27.

Schultz, W., Dayan, P., and Montague, P.R. (1997). A neural substrate of prediction and reward. *Science* *275*, 1593–1599.

Shoval, O., Goentoro, L., Hart, Y., Mayo, A., Sontag, E., and Alon, U. (2010). Fold-change detection and scalar symmetry of sensory input fields. *Proc. Natl. Acad. Sci.* *107*, 15995–16000.

Skinner, B.F. (2019). *The behavior of organisms: An experimental analysis* (BF Skinner Foundation).

Sola, J., and Sevilla, J. (1997). Importance of input data normalization for the application of neural networks to complex industrial problems. *IEEE Trans. Nucl. Sci.* *44*, 1464–1468.

Solomon, R.L. (1980). The opponent-process theory of acquired motivation: The costs of pleasure and the benefits of pain. *Am. Psychol.* *35*, 691–712.

Spanagel, R., Herz, A., and Shippenberg, T.S. (1992). Opposing tonically active endogenous opioid systems modulate the mesolimbic dopaminergic pathway. *Proc. Natl. Acad. Sci.* *89*, 2046–2050.

Stalder, T., Kirschbaum, C., Heinze, K., Steudte, S., Foley, P., Tietze, A., and Dettenborn, L. (2010). Use of hair cortisol analysis to detect hypercortisolism during active drinking phases in alcohol-dependent individuals. *Biol. Psychol.* *85*, 357–360.

Sutton, R.S., and Barto, A.G. (2018). *Reinforcement learning: An introduction* (MIT press).

Swann, H.G. (1940). THE PITUITARY-ADRENOCORTICAL RELATIONSHIP. *Physiol. Rev.* *20*, 493–521.

Szigethy, E., Conwell, Y., Forbes, N.T., Cox, C., and Caine, E.D. (1994). Adrenal weight and morphology in victims of completed suicide. *Biol. Psychiatry* *36*, 374–380.

Thorndike, E.L. (1927). The law of effect. *Am. J. Psychol.* *39*, 212–222.

Tobler, P.N., Fiorillo, C.D., and Schultz, W. (2005). Adaptive Coding of Reward Value by Dopamine Neurons. *Science* *307*, 1642–1645.

- Trigo, J.M., Zimmer, A., and Maldonado, R. (2009). Nicotine anxiogenic and rewarding effects are decreased in mice lacking  $\beta$ -endorphin. *Neuropharmacology* 56, 1147–1153.
- Tsigos, C., and Chrousos, G.P. (2002). Hypothalamic–pituitary–adrenal axis, neuroendocrine factors and stress. *J. Psychosom. Res.* 53, 865–871.
- Ulrich-Lai, Y.M., Figueiredo, H.F., Ostrander, M.M., Choi, D.C., Engeland, W.C., and Herman, J.P. (2006). Chronic stress induces adrenal hyperplasia and hypertrophy in a subregion-specific manner. *Am. J. Physiol.-Endocrinol. Metab.* 291, E965–E973.
- Van Ree, J.M. (1996). Endorphins and experimental addiction. *Alcohol* 13, 25–30.
- Vescovi, P.P., Coiro, V., Volpi, R., Giannini, A., and Passeri, M. (1992). Rapid communication: Plasma  $\beta$ -endorphin, but not met-enkephalin levels are abnormal in chronic alcoholics. *Alcohol* 27, 471–475.
- de Waele, J.-P., and Gianoulakis, C. (1993). Effects of Single and Repeated Exposures to Ethanol on Hypothalamic  $\beta$ -Endorphin and CRH Release by the C57BL/6 and DBA/2 Strains of Mice. *Neuroendocrinology* 57, 700–709.
- Westlund, K.N., Aguilera, G., and Childs, G.V. (1985). Quantification of Morphological Changes in Pituitary Corticotropes Produced by in Vivo Corticotropin-Releasing Factor Stimulation and Adrenalectomy. *Endocrinology* 116, 439–445.
- Winter, J., Bevan, S., and Campbell, E.A. (1995). Capsaicin and pain mechanisms. *Br. J. Anaesth.* 75, 157–168.
- Wise, R.A. (2004). Dopamine, learning and motivation. *Nat. Rev. Neurosci.* 5, 483–494.
- Wise, R.A., and Koob, G.F. (2014). The Development and Maintenance of Drug Addiction. *Neuropsychopharmacology* 39, 254–262.
- Yao, J., and Qin, F. (2009). Interaction with Phosphoinositides Confers Adaptation onto the TRPV1 Pain Receptor. *PLoS Biol.* 7.
- Young, E.A. (1989). Adrenalectomy increases  $\beta$ -Lipotropin secretion over  $\beta$ -endorphin secretion from anterior pituitary corticotrophs. *Life Sci.* 45, 2233–2237.
- Young, E., and Akil, H. (1985). Changes in releasability of ACTH and beta-endorphin with chronic stress. *Neuropeptides* 5, 545–548.